MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

THE DoD GATEWAY INFORMATION SYSTEM

October 1985

DTIC
ELECTE
DEC 1 1 1985
E

Office of Information Systems and Technology
DEFENSE TECHNICAL INFORMATION CENTER
Cameron Station
Alexandria, VA 22304-6145

85 12 11 028

AD-A161 701

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified/Unlimited	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
	Approved for public release; distribution
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
DTIC/TR-85/16	

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Defense Technical Information Center	DTIC	

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
Cameron Station Alexandria, VA 22304-6145	

8a. NAME OF FUNDING/SPONSORING ORGANIZATION Defense Technical Information Center	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
	DTIC	

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
Cameron Station Alexandria, VA 22304-6145	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
	658015			05000056

11 TITLE (Include Security Classification)

The DoD Gateway Information System

12. PERSONAL AUTHOR(S)
Cotter, Gladys A.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM _____ TO _____	8510	9

16 SUPPLEMENTARY NOTATION

Presented at the IEEE Conference, Williamsburg, VA

17	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Gateway, Front-End, Interfaces, Downloading, Networking,
5	2		Libraries, End-Users
9	2		

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The Defense Technical Information Center (DTIC) is sponsoring development of a
DoD Gateway Information System (DGIS) to provide online, streamlined methods for
identifying, accessing, searching and analyzing data from heterogenous databases
of interest to the DoD community. Present-day access to information resources
(data- bases) is limited since each database has its own complex access
procedure and command language. In addition, results from multiple databases
cannot be combined or analyzed easily by the user. The Gateway will provide DoD
researchers and managers access to many different databases using a single,
simple access procedure. Queries of databases will be performed using a common
command language, relieving the system user of the need to learn and master

(Continued on reverse)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified/Unlimited

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Gladys A. Cotter	(202) 274-7661	DTIC-EA

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified/Unlimited

19. ABSTRACT (Continued)

separate languages and procedures for each **database accessed**. A prototype
system is under development at the Lawrence **Livermore** National Laboratory. The
characteristics required in the DGIS and **the development** approach for designing
a prototype system are described.

THE DoD GATEWAY INFORMATION SYSTEM

by

Gladys A. Cotter

ABSTRACT

The Defense Technical Information Center (DTIC) is sponsoring development of a DoD Gateway Information System (DGIS) to provide online, streamlined methods for identifying, accessing, searching and analyzing data from heterogenous databases of interest to the DoD community. Present-day access to information resources (databases) is limited since each database has its own complex access procedure and command language. In addition, results from multiple databases cannot be combined or analyzed easily by the user. The Gateway will provide DoD researchers and managers access to many different databases using a single, simple access procedure. Queries of databases will be performed using a common command language, relieving the system user of the need to learn and master separate languages and procedures for each database accessed. A prototype system is under development at the Lawrence Livermore National Laboratory. The characteristics required in the DGIS and the development approach for designing a prototype system are described.

INTRODUCTION

The Defense Technical Information Center (DTIC) is charged with providing information services to the Department of Defense scientific and technical community. These services range from collecting and disseminating bibliographic information to sponsoring and directing research into innovative information handling technologies. Through this research, DTIC actively seeks ways to promote access to and utilization of, Scientific and Technical Information (STI) databases and online services and networks relevant to the conduct and management of research and engineering (R&E) programs. One of the most important efforts in this area is the development of a DoD Gateway Information System (DGIS). The DGIS will provide online, streamlined methods for identifying, accessing, searching, post-processing, and analyzing data from heterogeneous databases of interest to the DoD R&E community.

The necessity for the DGIS springs from the burgeoning proliferation of databases containing STI and the absence of accepted information handling standards within the industry. These factors of proliferation and lack of standards have produced severe barriers between information seekers (in our case, DoD researchers) and the information they require.

The current method for searching a database by use of a remote terminal requires that the researcher identify and access an appropriate distant computer and follow the unique search practices that have been programmed into it. The normal search requires that several databases be accessed, probably more than once each, and the researcher is burdened with interpreting and following a different instruction manual for every system. The product of the search is a

volume of printed matter that must be culled for the relevant material that is to be retained for use. For the infrequent user, most of the time and effort expended in a search are nonproductive; they are given over to identifying appropriate databases, accessing them, reading instruction manuals, and cutting and pasting printouts. The need is for the resulting information product, which takes relatively little time to assemble. The rest of the search process is expensive overhead.

The DGIS is being developed to eliminate the unproductive portion of the search process and allow researchers to spend their time utilizing the resultant information.

Our ultimate objective is to develop a system which can respond to a researcher's information need by locating the appropriate databases, conversing with them on the researcher's behalf, and providing a single, final, relevant information product. We are proceeding towards this goal in stages which we believe are realistic from technological and budgetary viewpoints.

DOD GATEWAY INFORMATION SYSTEM (DGIS) CHARACTERISTICS

The major requirements of the DGIS were established through DTIC user community surveys and site visits. In the process of identifying requirements for the DGIS, six critical areas surfaced. These areas are a gateway user interface, a directory of databases, database connection routines, common data retrieval routines, simultaneous search capabilities, and data analysis and post-processing routines.

Gateway User Interface

Designing a Gateway user interface so that the Gateway itself is simple to use, is the key to the success of this system. The DoD researcher may choose to interrogate the system directly or request that an intermediary such as a librarian or information specialist perform the search. The optimal interface for a DoD researcher or so-called "end user" of information is different from the optimal interface for an information intermediary. Therefore, distinctive system interfaces or user views must be tailored and designed for specific categories of users.

In addition, experience has taught us that even within the same user category different modes of operation are required. Novice and expert modes are vital and an intermediate mode is highly desirable. This allows an infrequent or new user to utilize numerous tutorial aids or menus to interact with the system at the novice level. (The frequent user also has the option of using these aids.)

The system becomes almost self-teaching and the notorious learning curve syndrome which often turns into a drop-out syndrome is avoided. As the user becomes proficient in the use of the system, these once helpful aids become cumbersome and repetitive. Frequent users, having learned the system, need to have an expert or intermediate mode available so that they can execute system options swiftly and directly.

Directory of Databases

There are over 3000 databases available online that contain scientific and technical information of interest to the DoD, and that number will continue to increase. The availability of these databases is a double-edged sword to the researcher. It presents the opportunity to acquire pertinent information, expand one's knowledge base, and accelerate the pace or increase the quality of the research. Conversely, the vast number of databases available makes maintaining an awareness of their existence and scope an awesome task for any researcher or even an information intermediary who has this responsibility.

To alleviate this problem, a Directory of Databases will form the core of the DGIS. The Directory will be maintained centrally by a staff at DTIC and shared throughout the DoD community so that a common body of knowledge concerning database existence will pervade the community. The directory itself will contain information on the content and scope of the available databases. The Directory will be subject-searchable so that upon entering the topic of interest the researcher will be provided with a list of appropriate databases.

Database Connection Routines

It is difficult, time consuming, and often frustrating for infrequent users to master a plethora of database "sign-on" procedures. Such procedures include dialing telephone numbers to access database host computers directly or to access valued-added networks (VANs) such as TYMNET and TELENET which will route the user to the database host. They also include executing system-specific protocols and entering appropriate password information. A single misplaced carriage return can bar a would-be user from connection. Even a seasoned user who requires access to more than five different database systems experiences difficulties which grow exponentially as the number of databases increases. Users who require access to multiple databases often resort to posting access methods, including passwords, in public areas. This defeats security, one of the basic reasons for these procedures in the first place.

Through the DGIS, database connection routines will be automated and protected. Authorized users will be able to issue a simple connect command and be linked with the information resource they desire to search.

Common Data Retrieval Routines

As noted earlier, the proliferation of databases has been accompanied by a lack of accepted design standards within the industry. This is particularly evident in the area of commands for information retrieval. Attempts to develop an International Standards Organization standard in this area have been unsuccessful to date. On the most basic level, the command to initiate a search varies from system to system. Neither the DoD researcher nor the information intermediary can afford the time to master the command language of each online system he or she accesses and keep up with their changes. In the case where particular databases are accessed infrequently, the expertise required to retrieve even moderately relevant information may take years to develop.

The DGIS will eventually support a common data retrieval routine for querying diverse databases. This feature will relieve the user of the need to learn and master separate commands and protocols for each database accessed.

The DGIS will be responsible for translating between its retrieval command set the native command sets of the diverse databases it queries on the user's behalf. All translation and protocol conversion will be transparent to the user.

Simultaneous Search Capabilities

Information relevant to the researcher's need will often be scattered among numerous databases. Therefore, the need exists to run the same search query against multiple databases simultaneously. Search results will be viewable on the terminal screen foreground, or may be relegated to the background freeing the screen for other activities. All search results will be downloaded from the remote database to the user's files on the DGIS as directed by the user and will be accessible by a single terminal.

Data Analysis and Post-Processing Routines

Often information retrieved from diverse databases requires analysis or post-processing to become useful to the researcher. Automated methods for analyzing the data are often the only practical way to deal with the sheer volume of information obtained. The ability to reformat, merge, sort and analyze data downloaded from remote databases to the DGIS catalyzes the transformation of an information glut into information gold.

DEVELOPMENT APPROACH

Having identified the major requirements of the DGIS, we sought to establish whether or not a software product was available which could serve as the basis for the DGIS. The software package which most closely met our needs and showed potential for being developed further to embody the DGIS characteristics was brought to our attention by the Department of Energy (DOE). DOE was sponsoring the Lawrence Livermore National Laboratory (LLNL) in its development of an intelligent gateway, the Technology Information System (TIS).

TIS was running in prototype mode at LLNL on a VAX 780 utilizing the UNIX operating system and the INGRES database management system. The communications and post-processing capabilities available on TIS were highly-applicable to the DoD requirements. Some of these features are highlighted below:

Communications Capabilities

Communications capabilities are the backbone of any gateway system and TIS has many outstanding features here. Users can access TIS via TYMNET, ARPANET, FTS, WATS and commercial phone lines. After login, many communications options are available. I will focus on electronic mail, write, link, connect, dial, and download.

Electronic Mail

Electronic mail service is available to all TIS users twenty-four hours a day. Standard electronic mail features such as send, receive, answer, and forward, are incorporated. Mail messages can be sent simultaneously to multiple addresses, with lengthy documents attached if needed. Users recognize the benefits of being able to communicate with numbers of people at the same time

and of avoiding the call-back routine. Messages can be filed for future
reference or deleted from the system upon command.

Write

Write is another communications option which allows users online to
communicate with each other via their terminals. You first enter the command
%WHO to get a display list of who is currently online. You then enter the
command %WRITE, followed by the name of the user you wish to communicate with,
which notifies that user, who then has the option of responding. The WRITE
command is only useful, of course, when parties who want to communicate are at
their terminals, by chance or arrangement, at the same time.

Link

The LINK command allows users at different and various locations to link
their terminals so that they are viewing the same data display. All users have
control over the display and can issue commands at will. Of course, linking
necessitates a cooperative spirit and some coordination.

Connect

The CONNECT command provides users with automatic access to information
resources. Users do not have to know telephone numbers, ARPANET locations,
passwords, access protocol or logout protocol. The user issues the CONNECT
command and a data resource name. TIS then attempts to establish a connection
to the resource and logs the user in. TIS uses TYMNET, TELENET, ARPANET,
COMMERCIAL TELEPHONE, and FTS to establish connections.

The CONNECT command can be used to access information centers worldwide.
In order to be eligible to use the CONNECT command for access to a resource, a
TIS user establishes an account with that resource and obtains the required
access identification information, such as passwords, to be programmed into the
gateway by the TIS Database Administrator. The billing process is unaffected by
gateway access. Vendors maintain the same billing structure and users maintain
the same reimbursement structure, regardless of the TIS access procedures. TIS
has several levels of security to ensure that password integrity is not
violated.

Dial

Users who wish to access a resource other than those listed in the TIS
resource directory take advantage of the DIAL command, rather than the CONNECT
command. DIAL allows users to call any information center, computer, or
terminal, no matter where the location. Using DIAL implies that the user knows
the necessary passwords and telephone numbers. DIAL allows the user to access
an off-network facility while retaining TIS capabilities such as downloading and
file transfer.

Downloading

Once you are connected to a resource through TIS, you can download data
from that resource. Downloading data opens many options to you. For example,
you can review it at your own pace, merge it with other data, and share it with

other users by allowing them to access your file. You can also transfer your file to other users so that they can manipulate the data to suit their own needs. TIS allows you to share your data selectively on a worldwide basis.

Post Processing

TIS offers a library of post-processing routines for numeric and bibliographic data. In order to execute post-processing routines, users must download the data into a TIS file. Post-processing routines for bibliographic data are available for selected resources. Some of the available routines are REVIEW, PLOT, PERMUTE, CROSS-CORRELATE, and CONCORD.

REVIEW allows users to process citations and determine relevance at their convenience. Users are presented with the author, title, date and several lines of an abstract. Based on this information they may choose to continue to work with the citation or discard it and move on to the next. If they continue to work with the citation, they may add local options, which include assigning relevancy values and index categories that are searchable. Users also can flag citations for which they wish to order the full text, plus add their own comments to a citation.

The PLOT routine allows users to generate bar charts representing the yearly publication rate for a subject area, personal author, or corporate author. This type of graphic representation makes growth trends immediately apparent.

PERMUTE provides statistics on the frequency of occurrence for descriptive terms in the citations. Single and compound expressions containing up to four terms are analyzed. These terms are presented in alphabetic order, preceded by the number of occurrences.

The CROSS-CORRELATION and CONCORD routines analyze the relationships among data elements chosen by the user. These routines provide intelligence that is very tedious to extract manually from standard bibliographies.

STATUS

Our review of existing TIS capabilities and discussions between the DOE and the DoD where mutual goals regarding information access were identified, led DTIC to enter into a joint DOD-DOE development effort. TIS is serving as the basis for, and LLNL is the major developer of, the DGIS.

At the present time, all DGIS development, test, and evaluation is taking place on the prototype system at LLNL. DTIC is sponsoring a number of DoD user entities who have agreed to test the system in their operations and make recommendations regarding its evolution into a DoD Gateway Information System. These users are issued passwords and dial into the LLNL prototype via TYMNET or WATS lines. TIS orientations are provided at DTIC or at the user's location through TIS linking technology.

Testing of the DGIS prototype will begin in October 1985 and will continue for a 12-month period. The purpose of the test is to demonstrate "proof-of-concept." To this end, the characteristics required in the DGIS will be tested within the limited universe of seven diverse database systems.

Through the DGIS prototype, users will be able to automatically connect to heterogenous databases, simultaneously search them and post-process and analyze the retrieved data. A directory for these database systems has been developed. The common data retrieval routines are still under development. A test version is scheduled for implementation on the prototype in April 1986. It is anticipated that artificial intelligence or expert system applications may prove very valuable in this area. Success of the prototype will result in its operational implementation and expansion of the number and type of databases accessible.

Implementation of the DGIS will provide DoD's scientific and technical community with a powerful, responsive information tool. The DGIS will render timely, comprehensive information to DoD research, development, and engineering programs. The productivity enhancement within the community resulting from this information will more than offset the investment made in DGIS development and operating cost.

END

FILMED

1-86

DTIC